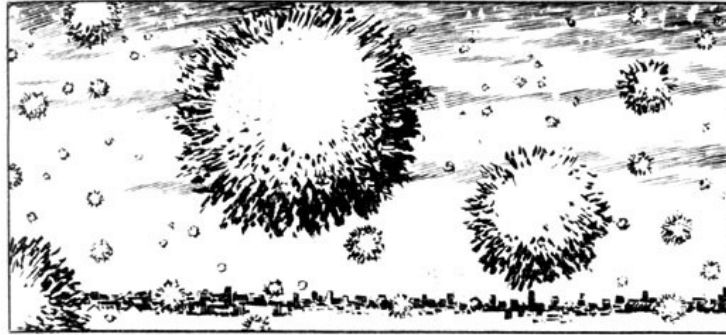


---

Álgebra II (Curso 23)  
Segundo cuatrimestre, 2021  
NOTAS EN LA EMERGENCIA SANITARIA:  
BORRADORES PARA LA CLASE DEL 6 DE DICIEMBRE  
Sebastian GRYNBERG

---



*El único héroe válido es el héroe “en grupo”,  
nunca el héroe individual, el héroe solo.*

H. G. OESTERHELD

ÍNDICE

1. Descomposición en valores singulares	2
1.1. Matrices diagonales	2
1.2. Bases ortonormales y diagonalización de matrices	2
1.3. Construcción	3
1.4. Proyecciones	4
1.5. DVS reducida	4
1.6. Seudoinversa de Moore-Penrose	4
2. El problema de mínimos cuadrados	5
2.1. La solución de norma mínima	5
2.2. Relación con la pseudoinversa de Moore-Penrose	7
3. Deformación de la esfera unitaria	7
3.1. Longitudes máximas y mínimas	7
3.2. Imagen de la esfera unitaria	8

## 1. DESCOMPOSICIÓN EN VALORES SINGULARES

## 1.1. Matrices diagonales.

Diremos que una matriz  $D \in \mathbb{R}^{m \times n}$  es diagonal si  $D_{ij} = 0$  para todo  $i \neq j$ . Si  $D \in \mathbb{R}^{m \times n}$  es diagonal usaremos la notación

$$D = \text{diag}_{m \times n}(d_1, \dots, d_q), \text{ donde } q = \min\{m, n\}$$

para indicar que  $D_{ii} = d_i$  para todo  $i \in \{1, \dots, q\}$ .

## 1.2. Bases ortonormales y diagonalización de matrices.

Una *descomposición en valores singulares* de  $A \in \mathbb{R}^{m \times n}$  (DVS) es una factorización de la matriz  $A$  en un producto de la forma  $U\Sigma V^T$ , donde

1.  $U \in \mathbb{R}^{m \times m}$  es una matriz ortogonal cuyas primeras  $r$  columnas  $u_1, \dots, u_r$  son una base ortonormal de  $\text{col}(A)$  y sus restantes  $m - r$  columnas  $u_{r+1}, \dots, u_m$  son una base ortonormal de  $\text{nul}(A^T)$ .
2.  $\Sigma \in \mathbb{R}^{m \times n}$  es una matriz diagonal de la forma  $\text{diag}_{m \times n}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ , donde  $r = \text{rango}(A)$  y  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ .
3.  $V \in \mathbb{R}^{n \times n}$  es una matriz ortogonal cuyas primeras  $r$  columnas  $v_1, \dots, v_r$  son una base ortonormal de  $\text{col}(A^T)$  y sus restantes  $n - r$  columnas  $v_{r+1}, \dots, v_n$  son una base ortonormal de  $\text{nul}(A)$ .

**Nomenclatura.** Los coeficientes  $\sigma_1, \sigma_2, \dots, \sigma_r$  se denominan *valores singulares* de  $A$ . Las columnas de  $U$ , *vectores singulares izquierdos* de  $A$  y las columnas de  $V$  *vectores singulares derechos* de  $A$ .

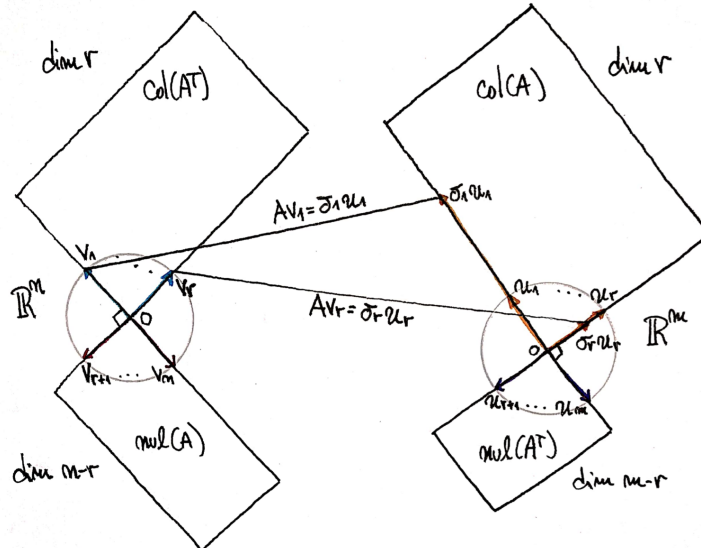


FIGURA 1. DVS: bases ortonormales que diagonalizan  $A$ . Nótese que si  $x = \sum_{j=1}^n \xi_j v_j$ , entonces  $Ax = \sum_{j=1}^r \sigma_j \xi_j u_j$ . En particular,  $\|Ax\|^2 = \sum_{j=1}^r \sigma_j^2 \xi_j^2$ .

**Nota Bene.** Nótese que una descomposición en valores singulares de  $A$  proporciona dos bases ortonormales: una de  $\mathbb{R}^n$  compuesta por las columnas de  $V$  y otra de  $\mathbb{R}^m$  compuesta por las columnas de  $U$ . Cuando se las utiliza para representar los vectores del dominio y la imagen de la transformación lineal  $T(x) = Ax$ , la naturaleza de la misma se vuelve transparente: simplemente dilata algunas componentes y contrae otras, de acuerdo con las magnitudes de los valores singulares, y posiblemente descarta componentes o agrega ceros, según sea necesario para tener en cuenta el cambio de dimensión.

### 1.3. Construcción.

**Paso 1.** La matriz  $V \in \mathbb{R}^{n \times n}$  se obtiene diagonalizando ortogonalmente la matriz simétrica  $A^T A \in \mathbb{R}^{n \times n}$ . La base ortonormal de autovectores de  $A^T A$  que compone las columnas de  $V = [v_1 \ v_2 \ \dots \ v_n]$  se ordena de manera tal que la matriz diagonal

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

que corresponde a la diagonalización ortogonal de  $A^T A$ ,

$$A^T A = V \Lambda V^T,$$

tenga los autovalores ordenados de mayor a menor:  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ . Como el rango de  $A$  es  $r$ , tenemos que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$  y  $\lambda_{r+1} = \dots = \lambda_n = 0$ . En particular, las primeras  $r$  columnas de  $V$  constituyen una base ortonormal de  $\text{col}(A^T)$  y las siguientes  $n - r$  columnas una base ortonormal de  $\text{nul}(A)$ .

**Paso 2.** Se calculan los valores singulares de  $A$ ,  $\sigma_j = \sqrt{\lambda_j}$ ,  $j \in \mathbb{I}_r$  y se utilizan para construir las primeras  $r$  columnas de la matriz  $U = [u_1 \ u_2 \ \dots \ u_m]$  de la siguiente manera

$$u_j = \frac{1}{\sigma_j} A v_j, \quad j \in \mathbb{I}_r.$$

El conjunto  $\{u_1, u_2, \dots, u_r\}$  constituye una base ortonormal de  $\text{col}(A)$ . Los  $m - r$  vectores restantes  $u_{r+1}, \dots, u_m$  se pueden construir aplicando el algoritmo de Gram-Schmidt a una base del espacio  $\text{nul}(A^T)$ .

**Paso 3.** La matriz  $\Sigma \in \mathbb{R}^{m \times n}$  se obtiene definiendo

$$\Sigma_r = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix},$$

y se completa con las matrices nulas que sean necesarias para contemplar las dimensiones  $m \times n$ :

$$\Sigma = \begin{bmatrix} \Sigma_r & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}.$$

#### 1.4. Proyecciones.

Consideramos una descomposición en valores singulares  $A = U\Sigma V^T$  de una matriz  $A \in \mathbb{R}^{m \times n}$  de rango  $r$ . Particionando  $U$  y  $V$  en la forma

$$U = [U_r \quad \tilde{U}_r], \quad V = [V_r \quad \tilde{V}_r],$$

donde

$$\begin{aligned} U_r &= [u_1 \quad \cdots \quad u_r] \in \mathbb{R}^{m \times r}, \quad \tilde{U}_r = [u_{r+1} \quad \cdots \quad u_n] \in \mathbb{R}^{m \times (m-r)}, \\ V_r &= [v_1 \quad \cdots \quad v_r] \in \mathbb{R}^{n \times r}, \quad \tilde{V}_r = [v_{r+1} \quad \cdots \quad v_n] \in \mathbb{R}^{n \times (n-r)}, \end{aligned}$$

tenemos

$$\begin{aligned} V_r V_r^T &= P_{\text{col}(A^T)}, & \tilde{V}_r \tilde{V}_r^T &= P_{\text{nul}(A)}, \\ U_r U_r^T &= P_{\text{col}(A)}, & \tilde{U}_r \tilde{U}_r^T &= P_{\text{nul}(A^T)}. \end{aligned}$$

*Demostración.* Recuérdese que si  $\{w_1, w_2, \dots, w_k\} \subset \mathbb{R}^p$  es una base ortonormal de un subespacio  $\mathbb{S}$ , entonces

$$\begin{aligned} P_{\mathbb{S}}(x) &= \sum_{j=1}^k \langle x, w_j \rangle w_j = \sum_{j=1}^k (w_j^T x) w_j = \underbrace{[w_1 \quad w_2 \quad \cdots \quad w_k]}_W \begin{bmatrix} w_1^T x \\ w_2^T x \\ \vdots \\ w_k^T x \end{bmatrix} \\ &= WW^T x. \end{aligned}$$

□

#### 1.5. DVS reducida.

Desde un punto de vista puramente algebraico cualquier fila o columna nula de la matriz  $\Sigma$  es superflua. Las mismas se pueden eliminar expresando el producto de matrices  $A = U\Sigma V^T$  mediante matrices en bloque

$$A = [U_r \quad \tilde{U}_r] \begin{bmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} V_r^T \\ \tilde{V}_r^T \end{bmatrix} = [U_r \Sigma_r \quad \mathbf{0}] \begin{bmatrix} V_r^T \\ \tilde{V}_r^T \end{bmatrix} = U_r \Sigma_r V_r^T.$$

Solamente las primeras  $r$  columnas de  $U$  y de  $V$  contribuyen al desarrollo de  $A$ :

$$(1) \quad A = U_r \Sigma_r V_r^T = [u_1 \quad \cdots \quad u_r] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix} = \sum_{j=1}^r \sigma_j u_j v_j^T.$$

La factorización de la matriz  $A$  que aparece en (1) se denomina *descomposición en valores singulares reducida de  $A$* .

#### 1.6. Seudoinversa de Moore-Penrose.

Si  $A = U_r \Sigma_r V_r^T$  es una descomposición en valores singulares reducida de  $A$ , la matriz

$$A^\dagger = V_r \Sigma_r^{-1} U_r^T,$$

se denomina *la pseudoinversa de Moore-Penrose de  $A$* .

**Nota Bene.** Nótese que valen las siguientes identidades

$$A^\dagger A = P_{\text{col}(A^T)} \quad \text{y} \quad AA^\dagger = P_{\text{col}(A)}.$$

Estas identidades son el motivo por el cual la matriz  $A^\dagger$  se denomina la *seudoinversa de Moore-Penrose de  $A$* . Si  $\text{rango}(A) = n = m$ , se tiene que  $A$  es invertible y en tal caso  $A^\dagger = A^{-1}$  porque  $P_{\text{col}(A^T)} = P_{\text{col}(A)} = I$ . Cuando  $\text{rango}(A) = n$ , la matriz  $A^T A$  es invertible y en tal caso  $A^\dagger = (A^T A)^{-1} A^T$ , lo que se comprueba facilmente y queda como ejercicio para el lector.

## 2. EL PROBLEMA DE MÍNIMOS CUADRADOS

Dada una matriz  $A \in \mathbb{R}^{m \times n}$  y un vector  $b \in \mathbb{R}^m$  se trata de hallar todos los  $x \in \mathbb{R}^n$  que minimizan la distancia  $\rho = \|b - Ax\|$ .

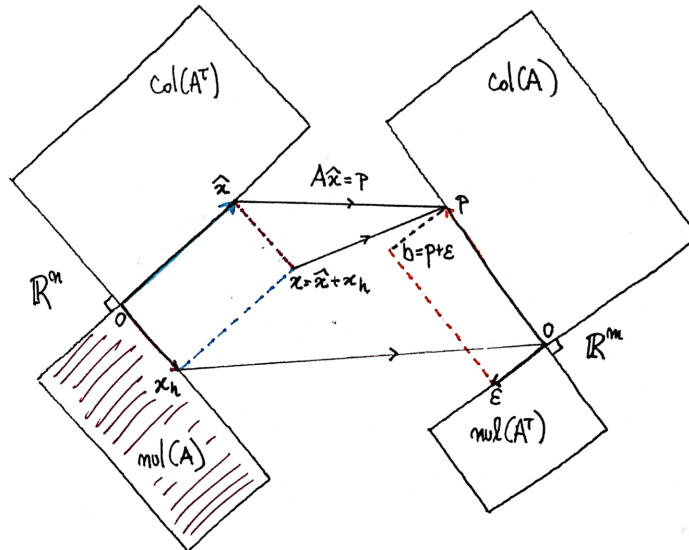


FIGURA 2. Una solución de mínimos cuadrados de  $Ax = b$  es cualquier vector  $x \in \mathbb{R}^n$  que minimiza  $\rho^2 = \|b - Ax\|^2$ , se obtienen resolviendo las ecuaciones normales  $A^T Ax = A^T b$ . La solución de mínimos cuadrados de norma mínima  $\hat{x}$  pertenece al subespacio  $\text{col}(A^T) = \text{nul}(A)^\perp$ , porque  $\|\hat{x} + x_h\|^2 = \|\hat{x}\|^2 + \|x_h\|^2 \geq \|\hat{x}\|^2$ .

### 2.1. La solución de norma mínima.

Sabemos que existe un único vector  $\hat{x} \in \text{col}(A^T)$  que minimiza  $\rho$  y que todos los demás se obtienen sumándole a  $\hat{x}$  los vectores del espacio  $\text{nul}(A)$ . En lo que sigue usaremos una DVS de  $A$  para determinar ese vector  $\hat{x}$ .

**Cambio de variables.** Si  $\text{rango}(A) = r$  y  $A = U\Sigma V^T$  es una DVS de  $A$  podemos escribir

$$(2) \quad b - Ax = b - U\Sigma V^T x = U(U^T b - \Sigma V^T x) = U(U^T b - \Sigma \xi),$$

donde  $\xi = V^T x$ .

Nótese que si  $V = [v_1 \ \cdots \ v_n]$ , entonces  $\xi$  es el vector de coordenadas de  $x$  respecto de la base ortonormal de  $\mathbb{R}^n$ ,  $\mathcal{B}_V = \{v_1, \dots, v_n\}$ , formada por las columnas de  $V$ . En efecto,

$$x = \sum_{j=1}^n \langle x, v_j \rangle v_j = \sum_{j=1}^n (v_j^T x) v_j = VV^T x.$$

Si  $U = [u_1 \ \cdots \ u_m]$  y  $\xi = [\xi_1 \ \cdots \ \xi_n]^T$ , el vector  $U^T b - \Sigma \xi$  tiene la forma

$$(3) \quad U^T b - \Sigma \xi = [u_1^T b - \sigma_1 \xi_1 \ \cdots \ u_r^T b - \sigma_r \xi_r \ u_{r+1}^T b \ \cdots \ u_m^T b].$$

Como  $U \in \mathbb{R}^{m \times m}$  es una matriz ortogonal su acción sobre  $\mathbb{R}^m$  es una isometría, lo que, junto a las identidades (2) y (3), nos permite calcular el error cuadrático

$$(4) \quad \rho^2 = \|U^T b - \Sigma \xi\|^2 = \sum_{j=1}^r (u_j^T b - \sigma_j \xi_j)^2 + \sum_{j=r+1}^m (u_j^T b)^2.$$

Lo único que es variable en la expresión del lado derecho de (4) es la suma

$$\sum_{j=1}^r (u_j^T b - \sigma_j \xi_j)^2$$

que depende de los valores de las primeras  $r$  coordenadas de  $x$  con respecto a la base  $\mathcal{B}_V$ :  $\xi_1, \dots, \xi_r$ . Como esa suma es una suma de cuadrados su mínimo valor se obtiene cuando, y solo cuando, todos los cuadrados valen 0. Eso significa que las coordenadas  $\xi_1, \dots, \xi_r$  son las soluciones de las ecuaciones

$$u_j^T b - \sigma_j \xi_j = 0, \quad j \in \mathbb{I}_r,$$

y por lo tanto,

$$\xi_j = \frac{u_j^T b}{\sigma_j}, \quad j \in \mathbb{I}_r.$$

Obtenemos así que los  $x \in \mathbb{R}^n$  que minimizan el error  $\rho = \|b - Ax\|$  son de la forma

$$x = \sum_{j=1}^r \frac{u_j^T b}{\sigma_j} v_j + \sum_{j=r+1}^n \xi_j v_j,$$

donde  $\xi_{r+1}, \dots, \xi_n \in \mathbb{R}$ . Esto, junto con el Teorema de Pitágoras, permite concluir que entre todos esos  $x$  el que tiene la norma mínima es

$$\hat{x} = \sum_{j=1}^r \frac{u_j^T b}{\sigma_j} v_j.$$

**Teorema 2.1.** Sean  $A \in \mathbb{R}^{m \times n}$  y  $b \in \mathbb{R}^m$ . Si  $\text{rango}(A) = r$  y  $A = U\Sigma V^T$  es una DVS de  $A$ , con  $U = [u_1 \ \cdots \ u_m]$  y  $V = [v_1 \ \cdots \ v_n]$ , entonces el vector

$$(5) \quad \hat{x} = \sum_{j=1}^r \frac{u_j^T b}{\sigma_j} v_j$$

minimiza el error cuadrático  $\|b - Ax\|^2$  y es el de norma mínima entre todos los vectores minimizadores del mencionado error. Además,

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|^2 = \|b - A\hat{x}\|^2 = \sum_{j=r+1}^m (u_j^T b)^2.$$

## 2.2. Relación con la pseudoinversa de Moore-Penrose.

**Nota Bene.** Nótese que, de acuerdo con el Teorema 2.1, la solución por mínimos cuadrados de  $Ax = b$  de longitud mínima es  $\hat{x} = A^\dagger b$ . Esto es así porque

$$\hat{x} = \sum_{j=1}^r \frac{u_j^T b}{\sigma_j} v_j = V_r \Sigma_r^{-1} U_r^T b.$$

En particular, la unicidad de  $\hat{x}$  implica que la pseudoinversa de Moore-Penrose no depende de la descomposición en valores singulares reducida de  $A$  que se utilice para construirla.

## 3. DEFORMACIÓN DE LA ESFERA UNITARIA

### 3.1. Longitudes máximas y mínimas.

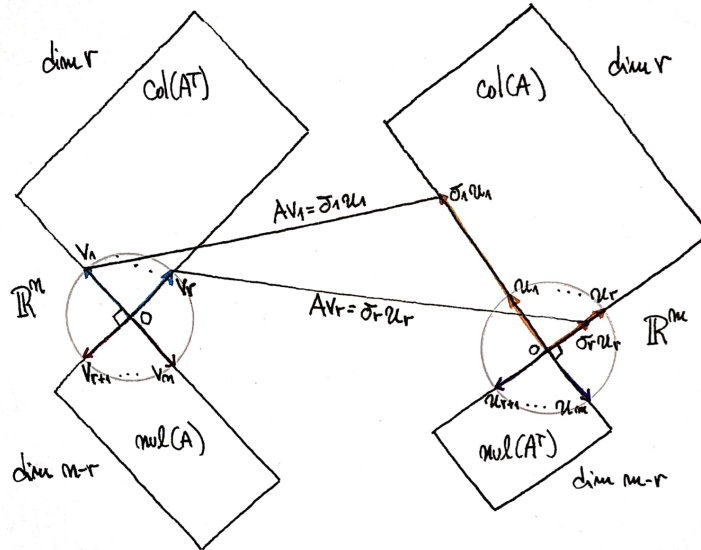


FIGURA 3. Esqueleto de la deformación de la esfera unitaria de  $\mathbb{R}^n$  bajo la acción de  $A$ :  $\|Av_j\| = \sigma_j$  para todo  $j \in \mathbb{I}_r$ , y  $\|Av_j\| = 0$  para todo  $j > r$ .

**Teorema 3.1.** Sea  $A = U_r \Sigma_r V_r^T$  una descomposición en valores singulares reducida de  $A$ , donde  $V_r = [v_1 \ \dots \ v_r]$ ,  $\Sigma_r = (\sigma_1, \dots, \sigma_r)$ . Vale que

1. Longitud máxima.

$$\max_{\|x\|=1} \|Ax\| = \sigma_1 \quad \text{y} \quad \operatorname{argmax}_{\|x\|=1} \|Ax\| = \left\{ x \in \mathbb{R}^n : x = \sum_{j=1}^k \xi_j v_j, \sum_{j=1}^k \xi_j^2 = 1 \right\},$$

donde  $k = \max \{j \in \mathbb{I}_r : \sigma_j = \sigma_1\}$ .

2. Longitud mínima.

a) Si  $r < n$ , entonces

$$\min_{\|x\|=1} \|Ax\| = 0 \quad \text{y} \quad \operatorname{argmin}_{\|x\|=1} \|Ax\| = \left\{ x \in \mathbb{R}^n : x = \sum_{j=r+1}^n \xi_j v_j, \sum_{j=r+1}^n \xi_j^2 = 1 \right\},$$

donde  $\{v_{r+1}, \dots, v_n\}$  es una base ortonormal de  $\operatorname{nul}(A)$ .

b) Si  $r = n$ , entonces

$$\min_{\|x\|=1} \|Ax\| = \sigma_n \quad \text{y} \quad \operatorname{argmin}_{\|x\|=1} \|Ax\| = \left\{ x \in \mathbb{R}^n : x = \sum_{j=k}^n \xi_j v_j, \sum_{j=k}^n \xi_j^2 = 1 \right\},$$

donde  $k = \min \{j \in \mathbb{I}_r : \sigma_j = \sigma_n\}$ .

*Demostración.* Descomponer  $x \in \mathbb{R}^n$  en la forma  $x = \sum_{j=1}^n \xi_j v_j$  y observar que

$$\|Ax\|^2 = \sum_{j=1}^r \sigma_j^2 \xi_j^2 \leq \sigma_1^2 \sum_{j=1}^r \xi_j^2 \leq \sigma_1^2 \|x\|^2.$$

Notar que si  $r = n$ , vale que

$$\|Ax\|^2 = \sum_{j=1}^n \sigma_j^2 \xi_j^2 \geq \sigma_n^2 \sum_{j=1}^n \xi_j^2 = \sigma_n^2 \|x\|^2.$$

Lo demás se cae de maduro. □

### 3.2. Imagen de la esfera unitaria.

Una manera de analizar cómo se deforma el espacio  $\mathbb{R}^n$  bajo la acción de la matriz  $A$  es considerar su acción sobre la esfera unitaria de  $\mathbb{R}^n$ ,

$$S_{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

Un elemento arbitrario  $x \in S_{n-1}$  se puede representar en la forma

$$x = \sum_{j=1}^n \xi_j v_j \quad \text{con} \quad \sum_{j=1}^n \xi_j^2 = 1.$$

En tal caso, la imagen de  $x$  por  $A$  es

$$Ax = \sum_{j=1}^r \sigma_j \xi_j u_j = \sum_{j=1}^r \eta_j u_j,$$

donde  $\eta_j = \sigma_j \xi_j$  para cada  $j \in \mathbb{I}_r$ . Como  $\frac{\eta_j}{\sigma_j} = \xi_j$ , tenemos que

$$\sum_{j=1}^r \frac{\eta_j^2}{\sigma_j^2} = \sum_{j=1}^r \xi_j^2 \leq \sum_{j=1}^n \xi_j^2 = 1.$$

Si  $A$  es de rango completo, i.e.,  $r = n$ , la desigualdad es una igualdad estricta. De lo contrario, faltan algunos  $\xi_j$  en la suma de la derecha, y la suma  $\sum_{i=1}^r \xi_j^2$  puede ser cualquier valor del intervalo  $[0, 1]$ . Esto muestra que  $A$  transforma la esfera unitaria de  $\mathbb{R}^n$  en un elipsoide de dimensión  $r$  con semiejes en las direcciones  $u_j$  y de longitudes  $\sigma_j$ . Cuando  $r = n$ , la imagen es la superficie del elipsoide, de lo contrario es el elipsoide sólido.

**Teorema 3.2.** Sea  $A \in \mathbb{R}^{m \times n}$  una matriz de rango  $r$ , y sea  $A = U\Sigma V^T$  una descomposición en valores singulares de  $A$ . Vale que

1. Si  $r = n$ , entonces

$$\{Ax : x \in \mathbb{S}_{n-1}\} = \left\{ y \in \mathbb{R}^m : y = \sum_{j=1}^n \eta_j u_j, \text{ con } \sum_{j=1}^n \frac{\eta_j^2}{\sigma_j^2} = 1 \right\}.$$

2. Si  $r < n$  tenemos que

$$\{Ax : x \in \mathbb{S}_{n-1}\} = \left\{ y \in \mathbb{R}^m : y = \sum_{j=1}^r \eta_j u_j, \sum_{j=1}^r \frac{\eta_j^2}{\sigma_j^2} \leq 1 \right\}.$$

**Nota Bene.** Nótese que, de acuerdo con lo anterior, el efecto de la acción de  $A$  sobre la esfera unitaria de  $\mathbb{R}^n$  se puede interpretar de la siguiente manera: primero colapsa  $n - r$  dimensiones de  $\mathbb{R}^n$ , luego distorsiona las dimensiones restantes, estirando y comprimiendo la esfera unitaria de dimensión  $r$  en un elipsoide, finalmente incrusta el elipsoide en  $\mathbb{R}^m$ .